Shrinking Horizon Model Predictive Control with Signal Temporal Logic Constraints under Stochastic Disturbances

Samira S. Farahani*, Rupak Majumdar**, Vinayak S. Prabhu†, Sadegh Soudjani‡

***Abstract*—We present shrinking horizon model predictive control for discrete-time linear systems under stochastic disturbances with constraints encoded as Signal Temporal Logic (STL) specification. The control objective is to satisfy a given STL specification with high probability against stochastic uncertainties while maximizing the robust satisfaction of an STL specification with minimum control effort. We formulate a general solution, which does not require precise knowledge of probability distributions of (possibly dependent) stochastic disturbances; only the bounded support of the density functions and moment intervals are used. For the specific case of disturbances that are normally distributed, we optimize the controllers by utilizing knowledge of the probability distribution of the disturbance. We show that in both cases, the control law can be obtained by solving optimization problems with linear constraints at each step. We experimentally demonstrate effectiveness of this approach by synthesizing a controller for an HVAC system.**

I. Introduction

We consider the control synthesis problem for stochastic discrete-time linear systems under path constraints that are expressed as temporal logic specifications and are written in signal temporal logic (STL) [21]. Our aim is to obtain a controller that robustly satisfies desired temporal properties with high probability despite stochastic disturbances, while optimizing additional control objectives. With focus on temporal properties defined on a finite path segment, we use model predictive control (MPC) scheme [3], [20] with a *shrinking horizon*: the horizon window is fixed and not shifted at each time step of the controller synthesis problem. We start with an initial prediction horizon dependent on the temporal logic constraints, compute the optimal control sequence for the horizon, apply the first step, observe the system evolution under the stochastic disturbance, and repeat the process (decreasing the prediction horizon by 1) till the end of the simulation time.

Our proposed setting requires solving three technical challenges in the MPC framework. First, in addition to optimizing control and state cost, the derived controller must ensure that the system evolution satisfies chance constraints arising from the STL specifications, i.e., closed-loop trajectories that depend on uncertain variables must satisfy specifications with high probability. Previous choices of control actions can impose temporal constraints on the rest of the path. The shrinking horizon approach guarantees that the previous actions will be taken into account when future control actions are computed. Second, for some temporal constraints, we may require that the system satisfies the constraints *robustly*: small changes to the inputs should not invalidate the temporal constraint. To ensure robust satisfaction, we use a quantitative notion of robustness for STL [9], [10]. We augment the control objective to maximize the expected robustness of an STL specification, in addition to minimizing control and state costs under chance constraints. Unfortunately, the resulting optimization problem is not convex. As a third contribution, we propose an approximation method for the solution of the optimization problem. We conservatively approximate chance constraints by linear inequalities and compute an upper bound for the expected value of the robustness function that appears in the objective function.

Recently receding horizon control with STL constraints has been studied in [12], [24], [26], where the worst-case MPC optimization problem is solved by assuming disturbances taking values from a bounded polytope. An overview of stochastic control under chance constraints can be found in [28] and customized approaches for normally distributed uncertainties are presented in [29]. Chance-constrained MPC for deterministic systems with measurement noise has been addressed in [27]. It is also applied to drinking water networks [14] and to urban autonomous driving [6]. The work [25] addresses optimizations with constraints encoded via convex fragment of a logic known as PrSTL. The class of C2TL specifications is defined in [17] for deterministic systems, where the uncertainty is introduced only in the coefficients of atomic predicates.

In this paper, we assume in the general case that the disturbance has an arbitrary probability distribution with bounded domain and that we only know its support and first moment interval. In order to solve the optimization problem efficiently, we transform chance constraints into linear constraints. To this end, we employ concentration of measure inequalities [5] to conservatively approximate the feasible domain of the optimization specified by chance constraints. We also approximate the expected value of the robustness function using the moment interval of the disturbance to prevent numerical integration. For the special case where the disturbance is normally disturbed, we apply additional computational techniques. Clearly, the assumption of bounded support is not valid for this case. Instead of truncating normal distribution to obtain a bounded support, we employ a different approach based on quantiles of normally distributed random variables to replace chance constraints by linear constraints. We show that in this case, the expected value of the robustness function can be upper bounded based on techniques from [13] developed for approximating the expected value of max-affine expressions.

* Delft University of Technology, The Netherlands
s.farahani@tudelft.nl
** Max Planck Institute for Software Systems, Germany
rupak@mpi-sws.org
† Colorado State University, USA
vinayak.prabhu@colostate.edu
‡School of Computing, Newcastle University, United Kingdom
sadegh.soudjani@ncl.ac.uk
This research was supported in part by FCT grant SFRHBPD902672012

Our work extends the results of [11], where only the case of normal distribution is discussed, and gives more compact and efficient representations for transforming probabilistic constraints into linear constraints. We demonstrate the effectiveness of our approach by synthesizing a controller for a Heating, Ventilation and Air Conditioning (HVAC) system.

Notation. We use $\mathbb{R}$ for the set of reals and $\mathbb{N}:=\{0,1,2,\ldots\}$ for the set of non-negative integers. For $v\in\mathbb{R}^s$, its components are denoted by $v_k, k\in\{1,\ldots,s\}$. We use small letter y to indicate observations of a random vector Y. For a random variable X with values in $\mathbb{R}^n$ and probability distribution Pr, its support is defined as the smallest closed set C such that $\Pr[X\in C]=1$. We denote the support of X by I_X and its first moment by $\mathbb{E}[X]$.

II. Discrete-Time Stochastic Linear Systems

We consider time-variant discrete-time stochastic systems modeled by the difference equation

$$X(t+1)=A(t)X(t)+B(t)u(t)+W(t),\ X(0)=x_0, \tag{1}$$

where $X(t)\in\mathbb{R}^n, u(t)\in\mathbb{R}^m$, and $W(t)\in\mathbb{R}^n$ denote respectively the state, control input, and disturbance of the system at time instant t. Matrices $A(\cdot)\in\mathbb{R}^{n\times n}$ and $B(\cdot)\in\mathbb{R}^{n\times m}$ are possibly time-dependent system's matrices, and the initial state $X(0)$ is known. We assume that $W(0),\ldots,W(t)$ are mutually independent random vectors for all time instants t. We conduct our study for two cases: a) the disturbance signal has an arbitrary probability distribution with a bounded domain for which we only know the support and their first moment intervals; and b) the disturbance signal has a normal distribution. For any $t\in\mathbb{N}$, the state-space model (1) provides the following explicit form for $X(\tau)$, $\tau\geq t$, as a function of $X(t)$, $u(\cdot)$, and $W(\cdot)$,

$$X(\tau)=\Phi(\tau,t)X(t)+\sum_{k=t}^{\tau-1}\Phi(\tau,k+1)\left(B(k)u(k)+W(k)\right), \tag{2}$$

where $\Phi(\cdot,\cdot)$ is the *state transition matrix* of (1), defined as

$$\Phi(\tau,t)=\begin{cases}A(\tau-1)A(\tau-2)\ldots A(t) & \tau>t\geq 0\\ \mathbb{I}_n & \tau=t\geq 0,\end{cases}$$

with $\mathbb{I}_n$ being the identity matrix.

For a fixed positive integer N, and a given $t\in\mathbb{N}$, let $\tilde{u}(t:N)=[u^T(t),u^T(t+1),\ldots,u^T(N-1)]$ (vector $\tilde{W}(t:N)$ is defined similarly). Given system (1), and a time interval $[t:N]$, a (discrete-time) *stochastic process* can be defined as $\Xi(t:N)=X(t)X(t+1)\ldots X(N)$, corresponding to a finite sequence of random state variables. As the process $\Xi(t:N)$ depends on $X(t),\tilde{u}(t:N)$, and $\tilde{W}(t:N)$, we can rewrite $\Xi(t:N)$ in a more elaborative notation as $\Xi_N(X(t),\tilde{u}(t:N),\tilde{W}(t:N))$. Analogously, we define an *unbounded-time stochastic process* $\Xi=X(t)X(t+1)X(t+2)\ldots$, corresponding to an infinite sequence of random state variables.

III. Signal Temporal Logic

An infinite run of system (1) can be considered as a signal $\xi=x(0)x(1)x(2)\ldots$, which is a sequence of observed states. We consider Signal temporal logic (STL) formulas with bounded-time temporal operators defined recursively according to the grammar [21]: $\varphi ::= \top \mid \pi \mid \neg\varphi \mid \varphi\wedge\psi \mid \varphi\mathcal{U}_{[a,b]}\psi$; where $\top$ is the *true* predicate; π is a predicate of the form $\pi=\{\alpha(x)\geq 0\}$ with $\alpha:\mathbb{R}^n\to\mathbb{R}$ being an affine function of state variables; ψ is an STL formula; $\neg$ and $\wedge$ indicate negation and conjunction of formulas; and $\mathcal{U}_{[a,b]}$ is the *until* operator with $a,b\in\mathbb{R}_{\geq 0}$ and $a\leq b$.

A run ξ satisfies φ at time t, denoted by $(\xi,t)\models\varphi$, if the sequence $x(t)x(t+1)\ldots$ satisfies φ. Accordingly, ξ satisfies φ, denoted by $\xi\models\varphi$, if $(\xi,0)\models\varphi$. Semantics of STL formulas are defined as follows. Every run satisfies $\top$. For a run $\xi=x(0)x(1)x(2)\ldots$ and a predicate $\pi=\{\alpha(x)\geq 0\}$, we have $(\xi,t)\models\pi$ if $\alpha(x(t))\geq 0$. The run ξ satisfies $\neg\varphi$ if it does not satisfy φ; it satisfies $\varphi\wedge\psi$ if both φ and ψ hold. Finally, $(\xi,t)\models\varphi\mathcal{U}_{[a,b]}\psi$ if φ holds at every time step starting from time t before ψ holds, and additionally ψ holds at some time instant between $a+t$ and $b+t$. Moreover, we derive the other standard operators as follows. *Disjunction* $\varphi\vee\psi:=\neg(\neg\varphi\wedge\neg\psi)$, *eventually* as $\diamondsuit_{[a,b]}\varphi:=\top\mathcal{U}_{[a,b]}\varphi$, and *always* as $\square_{[a,b]}\varphi:=\neg\diamondsuit_{[a,b]}\neg\varphi$. For an unbounded-time stochastic process $\Xi=X(t),X(t+1),X(t+2),\ldots$, we denote by $\Pr(\Xi\models\varphi)$ the probability measure of the set of instantiations ξ of Ξ such that $\xi\models\varphi$.

Formula Horizon. The *horizon* of an STL formula φ, denoted by Δ, is the smallest $n\in\mathbb{N}$ such that the following holds for all signals $\xi=x(0)x(1)x(2)\ldots$ and $\xi'=x'(0)x'(1)x'(2)\ldots$: if

$$x(t+i)=x'(t+i)\ \forall i\in\{0,\ldots,n\}\Rightarrow(\xi,t)\models\varphi \text{ iff } (\xi',t)\models\varphi.$$

Thus, in order to determine whether a signal ξ satisfies an STL formula φ, we can restrict our attention to the signal prefix $x(0),\ldots,x(\Delta)$. This horizon can be upper-approximated by a bound, defined as the maximum over the sums of all nested upper bounds on the temporal operators, denoted by $\mathrm{len}(\varphi)$. Formally, $\mathrm{len}(\varphi)$ is defined recursively as:

$$\begin{aligned}&\mathrm{len}(\top)=\mathrm{len}(\pi)=0,\quad \mathrm{len}(\neg\varphi)=\mathrm{len}(\varphi),\\&\mathrm{len}(\varphi_1\wedge\varphi_2)=\max(\mathrm{len}(\varphi_1),\mathrm{len}(\varphi_2)),\\&\mathrm{len}(\varphi_1\ \mathcal{U}_{[a,b]}\ \varphi_2)=b+\max(\mathrm{len}(\varphi_1),\mathrm{len}(\varphi_2)).\end{aligned}$$

For example, for $\varphi=\square_{[0,4]}\diamondsuit_{[3,6]}\pi$, we have $\mathrm{len}(\varphi)=4+6=10$. For any STL formula φ, it is possible to verify that $\xi\models\varphi$ using only the finite run $x(0)x(1)\ldots x(\mathrm{len}(\varphi))$.

STL Robustness. In contrast to the above Boolean semantics, the quantitative semantics of STL [9], [18] assigns to each formula φ a real-valued function ρ^φ of signal ξ and t such that $(\xi,t)\models\varphi$ if $\rho^\varphi(\xi,t)>0$, and is defined recursively as

$$\begin{aligned}&\rho^\top(\xi,t)=+\infty,\qquad \rho^\pi(\xi,t)=\alpha(x(t)) \text{ with } \pi=\{\alpha(x)\geq 0\},\\&\rho^{\neg\varphi}(\xi,t)=-\rho^\varphi(\xi,t),\ \ \rho^{\varphi\wedge\psi}(\xi,t)=\min(\rho^\varphi(\xi,t),\rho^\psi(\xi,t)),\\&\rho^{\varphi\,\mathcal{U}_{[a,b]}\psi}(\xi,t)=\max_{i\in[a,b]}\left(\min(\rho^\psi(\xi,t+i),\min_{j\in[0,i)}\rho^\varphi(\xi,t+j))\right),\end{aligned}$$

where $x(t)$ refers to signal ξ at time t. Robustness of $\diamondsuit_{[a,b]}\varphi$ can be derived as $\rho^{\diamondsuit_{[a,b]}\varphi}(\xi,t)=\max_{i\in[a,b]}\rho^\varphi(\xi,t+i)$. Similarly, $\rho^{\square_{[a,b]}\varphi}(\xi,t)=\min_{i\in[a,b]}\rho^\varphi(\xi,t+i)$.

STL Robustness for Stochastic Processes. Analogous to robustness for signals $\xi=x(0)x(1)x(2)\ldots$, we define the *stochastic robustness* $\rho^{\varphi}(\Xi, t)$ of a formula φ (with bounded-time temporal operators) at time t with respect to the stochastic process Ξ, by replacing the concrete states $x(t)$ with the random state variables $X(t)$. It can be shown that the bounded-time stochastic process $\Xi(t:t+N) = X(t)X(1)\ldots X(t+N)$ with $N = \text{len}(\varphi)$ is sufficient to study the probabilistic properties of Ξ with respect to φ. Note that $\rho^{\varphi}(\Xi(t:t+N), t)$ is a random variable since affine operators, maximization, and minimization are measurable functions. We can also show that for any formula φ and constant $\delta \in (0,1)$, the stochastic process $\Xi = X(0)X(1)X(2)\ldots$ satisfies φ with probability $\geq 1-\delta$ (*i.e.* $\Pr(\Xi \models \varphi) \geq 1-\delta$) if $\Pr[\rho^{\varphi}(\Xi(0:N), 0) > 0] \geq 1-\delta$ for some $N \geq \text{len}(\varphi)$.

IV. Problem Statement

For system (1) with a given initial state $X(0) = x_0$, STL formulas φ and ψ, and a constant $N \geq \max(\text{len}(\varphi), \text{len}(\psi))$, the control problem can be defined as finding an optimal input sequence $\tilde{u}^*(0:N) = [u^{*T}(0), \ldots, u^{*T}(N-1)]$, that minimizes the expected value of a given objective function $J(\tilde{X}(0:N+1), \tilde{u}(0:N))$ subject to constraints on states and input variables, where $\tilde{X}(0:N+1) = [X^T(0), X(1)^T x, \ldots, X^T(N)]$. This optimization problem is defined as

$$\min_{\tilde{u}(0:N)} \ \mathbb{E}\left[J(\tilde{X}(0:N+1), \tilde{u}(0:N))\right] \quad \text{s.t.} \tag{3a}$$

$$X(t) = \Phi(t,0)x_0 + \sum_{k=0}^{t-1} \Phi(t,k+1)\left(B(k)u(k) + W(k)\right), \tag{3b}$$

$$\Pr\left[\Xi_N(x_0, \tilde{u}(0:N), \tilde{W}(0:N)) \models \varphi\right] \geq 1-\delta, \tag{3c}$$

$$\tilde{u}(0:N) \in U^N, \tag{3d}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator and the set $U \subset \mathbb{R}^m$ specifies the constraint set for the input variables. The chance constraints (3c) state that for a given $\delta \in (0,1)$, stochastic process Ξ_N should satisfy φ with probability $\geq 1-\delta$.

We consider the following objective function

$$J(\tilde{X}(0:N+1), \tilde{u}(0:N)) := J_{\text{robust}}(\tilde{X}(0:N+1)) + J_{\text{in}}(\tilde{u}(0:N)),$$

where the first term $J_{\text{robust}}(\tilde{X}(0:N+1)) := -\rho^{\psi}(\tilde{X}(0:N+1), 0)$ represents the negative value of the robustness function on STL formula ψ at time 0 that needs to be minimized; and the second term $J_{\text{in}}(\tilde{u}(0:N))$ reflects the cost on input variables defined based on infinity norm, one norm, or any piecewise constant function.

Remark 1: The above problem formulation enables us to distinguish the following two cases: we put the robustness of a formula in the objective function if the system is required to be robust with respect to satisfying the formula; we encode the formula in the probabilistic constraint if only satisfaction of the formula is important.

V. Model Predictive Control

Optimization problem (3) is an open-loop optimization that does not incorporate any information related to the observed states of the system. In order to include such information in the computation of the control input, instead of solving (3), we employ *shrinking horizon model predictive control* (SHMPC), which is summarized as follows: at time step one, we obtain a sequence of control inputs with length N (the prediction horizon) to optimize the cost function; then we only apply the first component of the obtained control sequence to the system and observe the next state; in the next time step, we fix the first component of the control sequence by its optimal value and hence, we only optimize for a control sequence of length $N-1$. As such, at each time step, the size of the control sequence decreases by 1.

A natural choice for the prediction horizon N in this setting with STL specifications φ and ψ in constraints and in the objective function is to set it greater than or equal to the bounds of the formulas, i.e., $N \geq \max(\text{len}(\varphi), \text{len}(\psi))$, with the length of formula defined in the previous section. This choice provides a conservative trajectory length required to make a decision about the satisfiability of the formula.

Let $\bar{X}(0:t:N+1) = [x^T(0), \ldots, x^T(t), X^T(t+1), \ldots, X^T(N)]$ where $x(0), \ldots, x(t)$ are the observed states up to time t and $X(\tau)$ is the random state variable at time $\tau > t$, also let $\bar{W}(0:t-1:N) = [w^T(0), \ldots, w^T(t-1), W^T(t), W^T(t+1), \ldots, W^T(N-1)]$ such that $w(0), \ldots, w(t-1)$ are disturbance realizations up to time $t-1$ and $W(\tau)$ are random disturbances at time $t \leq \tau \leq N-1$. Define $\bar{u}(0:t-1:N) = [u^{*T}(0), \ldots, u^{*T}(t-1), u^T(t), \ldots, u^T(N-1)]$ to be the vector of input variables such that $u^*(0), \ldots, u^*(t-1)$ are the obtained optimal control inputs up to time $t-1$ and $u(t), \ldots, u(N-1)$ are the input variables that need to be determined at time t.

Given formulas φ, ψ, observed states $x(0), x(1), \ldots, x(t)$, and obtained control inputs $u^*(0), \ldots, u^*(t-1)$ for system (1), the *stochastic SHMPC* optimization problem minimizes the expectation of cost function $J(\bar{X}(0:t:N+1), \bar{u}(0:t-1:N)) = J_{\text{robust}}(\bar{X}(0:t:N+1)) + J_{\text{in}}(\bar{u}(0:t-1:N))$, at each time instant $0 \leq t < N$, as

$$\min_{\tilde{u}(t:N)} \ \mathbb{E}\left[J(\bar{X}(0:t:N+1), \bar{u}(0:t-1:N))\right] \quad \text{s.t.} \tag{4a}$$

$$X(\tau) = \Phi(\tau,t)x(t) + \sum_{k=t}^{\tau-1} \Phi(\tau,k+1)\left(B(k)u(k) + W(k)\right), \tag{4b}$$

$$\Pr[\Xi_N(x_0, \bar{u}(0:t-1:N), \bar{W}(0:t-1:N)) \models \varphi] \geq 1-\delta \tag{4c}$$

$$\tilde{u}(t:N) \in U^{N-t}, \tag{4d}$$

where the expected value $\mathbb{E}[\cdot]$ in (4a) is conditioned on observed states $x(0), \ldots, x(t)$. Optimization variables in (4) are control inputs $\tilde{u}(t:N) = [u^T(t), \ldots, u^T(N-1)]$. We indicate the argument of minimum by $\tilde{u}_{opt}(t:N) = [u^T_{opt}(t), \ldots, u^T_{opt}(N-1)]$.

The following theorem states that by using SHMPC that keeps track of control inputs and observed states, the closed-loop system satisfies the STL specification φ with probability greater than or equal to $1-\delta$.

Theorem 2: Given $\delta \in (0,1)$ and STL formula φ, if the optimization problem (4) is feasible for all $t < N$, the computed optimal control sequence $\tilde{u}^*(0:N) = [u^{*T}(0), \ldots, u^{*T}(N-1)]$ ensures that the closed-loop system satisfies φ with probability greater than or equal to $1-\delta$.

Remark 3: Note that for having the result of Theorem 2, we only need feasibility at the last time step $t = N-1$. In practice, the optimization problem (4) might be infeasible for some $t < N$ due to the stochastic nature of the disturbance. Therefore, we should guide the optimization towards its feasible domain, which can be done by replacing φ with its relaxed version [26] and try to minimize the violation of constraints. Alternatively, for any $t<N$ that (4) is infeasible, we opt for maximizing the expectation of robustness of φ, $\mathbb{E}[\rho^{\varphi}(\bar{X}(0:,t,N+1),0)]$, without any chance constraint in order to obtain an input that is most likely to result in satisfaction of φ.

Computing the solution of optimization problem (4) requires addressing two main challenges: a) the expected value of the objective function (4a) is in general difficult to be calculated analytically as a function of $\bar{u}(t:N)$; b) it is hard to characterize the exact feasible set of the optimization restricted by the chance constraint (4c). We propose approximation methods in Sections VI–VII to respectively address these two challenges.

VI. Approximating the Objective Function

To solve the optimization problem (4), one needs to calculate the expected value of the objective function. One way to do this is via numerical integration methods [7], which is in general both cumbersome and time-consuming. In this section, we discuss an efficient method that computes an upper bound for the expected value of the objective function and we minimize this upper bound instead. We discuss computation of such upper bounds for both cases of disturbances with arbitrary probability distribution and with normal distribution. We first provide a canonical form for the STL robustness function which is the min-max or max-min of random variables. This result is inspired by [8], in which the authors provide such canonical forms for max-min-plus-scaling functions.

Theorem 4: For a given STL formula ψ, the robustness function $\rho^{\psi}(\Xi(0:N),0)$, and hence the function $J_{\text{robust}}(\bar{X}(0:t:N))$, can be written into a max-min canonical form

$$J_{\text{robust}}(\bar{X}(0:t:N)) = \max_{i\in\{1,\ldots,L\}} \min_{j\in\{1,\ldots,m_i\}} \left\{\eta_{ij} + \bar{W}(0:t:N)\lambda_{ij}\right\}, \tag{5}$$

and into a min-max canonical form

$$J_{\text{robust}}(\bar{X}(0:t:N)) = \min_{i\in\{1,\ldots,K\}} \max_{j\in\{1,\ldots,n_i\}} \left\{\zeta_{ij} + \bar{W}(0:t:N)\gamma_{ij}\right\}, \tag{6}$$

for some integers $K, L, n_1, \ldots, n_K, m_1, \ldots, m_L$, where λ_{ij} and γ_{ij} are column vectors as weights and η_{ij} and ζ_{ij} are affine functions of $\bar{u}(0:t:N)$ and x_0.

Remark 5: Note that any of the canonical forms (5) and (6) can be transformed to the other one utilizing identities

$$\begin{aligned}&\min\left(\max(f_1,f_2),\max(g_1,g_2)\right) = \\ &\quad\max\left(\min(f_1,g_1),\min(f_1,g_2),\min(f_2,g_1),\min(f_2,g_2)\right).\end{aligned}$$

and $-\max(f_1,f_2) = \min(-f_1,-f_2)$, for any f_1, f_2, g_1, and g_2.

Proof: Proof is inductive on the structure of ψ. Since (5) and (6) can be transformed to each other using identities of Remark 5, it is sufficient to work with and establish only one of them. The canonical forms are valid for any atomic predicate $\{\alpha(x)\geq 0\}$ evaluated at time τ. To see this, take the affine function $\alpha(x) := \alpha_0 + \alpha_1^T x$ and use state equation (2) to get $L = K = m_i = n_i = 1$ and $\zeta_{11} = \eta_{11}$ with

$$\eta_{11} := -\alpha_0 - \alpha_1^T\Phi(\tau,0)x(0) - \sum_{k=0}^{\tau-1}\alpha_1^T\Phi(\tau,k+1)B(k)u(k).$$

We also have $\gamma_{11} = \lambda_{11} = [\lambda_{11}^0;\lambda_{11}^1;\ldots;\lambda_{11}^{N-1}]$ with $\lambda_{11}^k = -\Phi^T(\tau,k+1)\alpha_1$ for $k<\tau$ and zero, otherwise. Suppose ρ^{ψ} has the form (5) with η_{ij} and λ_{ij}. Then $\rho^{\neg\psi} = -\rho^{\psi}$ has the form (6) with $\zeta_{ij} = -\eta_{ij}$, $\gamma_{ij} = -\lambda_{ij}$, and the same set of indices. If ρ^{ψ_1} and ρ^{ψ_2} have the canonical form (5) with $L^1, L^2, \eta_{ij}^1, \lambda_{ij}^2$, then $\rho^{\psi_1\vee\psi_2} = \max(\rho^{\psi_1},\rho^{\psi_2})$ will also have the form (5) with $L = L_1 + L_2$ and

$$\begin{cases}\eta_{ij} = \eta_{ij}^1 \text{ and } \lambda_{ij} = \lambda_{ij}^1 \text{ for } 1\leq i\leq L_1,\\ \eta_{ij} = \eta_{ij}^2 \text{ and } \lambda_{ij} = \lambda_{ij}^2 \text{ for } L_1+1\leq i\leq L_2.\end{cases}$$

Similar equalities hold for $\psi = \psi_1\wedge\psi_2$ but using canonical form (6). The same reasoning can be applied to $\psi_1\mathcal{U}\psi_2$. ■

A. Arbitrary probability distributions with bounded support

Suppose the elements of the stochastic vector $W(t)$, i.e., $W_k(t)$, $k\in\{1,\ldots,n\}$ have arbitrary probability distribution with known bounded support $I_{W_k(t)} = [a_k,b_k]$ and their first moments $\mathbb{E}[W_k(t)]$ belongs to the intervals $\mathbb{M}_{W_k(t)} = [c_k,d_k]$, with known quantities $a_k,b_k,c_k,d_k\in\mathbb{R}$. We denote by $I_{W(t)}$ and $\mathbb{M}_{W(t)}$ respectively as the product of intervals $I_{W_k(t)}$ and $\mathbb{M}_{W_k(t)}$, $k\in\{1,\ldots,n\}$. Under this assumption, the explicit form of $X(\cdot)$ in (2) implies that, for the observed value of $X(t)$ as $x(t)$, $X(\tau)$ is a random vector with the following interval of support and the first moment interval

$$I_{X(\tau)} = [\bar{a}_\tau + \bar{C}_\tau, \bar{b}_\tau + \bar{C}_\tau],\ \mathbb{M}_{X(\tau)} = [\bar{c}_\tau + \bar{C}_\tau, \bar{d}_\tau + \bar{C}_\tau] \tag{7}$$

where $\bar{C}_\tau = \Phi(\tau,t)x(t) + \sum_{k=t}^{\tau-1}\Phi(\tau,k+1)B(k)u(k)$, $[\bar{a}_\tau,\bar{b}_\tau] = \sum_{k=t}^{\tau-1}\Phi(\tau,k+1)I_{W(k)}$, and $[\bar{c}_\tau,\bar{d}_\tau] = \sum_{k=t}^{\tau-1}\Phi(\tau,k+1)\mathbb{M}_{W(k)}$. The elements of $\bar{a}_\tau,\bar{b}_\tau,\bar{c}_\tau$ and $\bar{d}_\tau$ are computed using the following operations on intervals extended naturally to vectors and matrix multiplications: for two arbitrary intervals $[a,b]$ and $[c,d]$, and constant $\lambda\in\mathbb{R}$, we have $[a,b]+[c,d] = [a+c,b+d]$ and $\lambda\cdot[a,b] = [\min(\lambda a,\lambda b),\max(\lambda a,\lambda b)]$.

The expected value of the objective function in (4) can be written as $\mathbb{E}[J_{\text{robust}}(\bar{X}(0:t:N+1))] + J_{\text{in}}(\bar{u}(0:t-1:N)))$, where $\bar{X}(0:t:N+1)$ includes un-observed states after t. The next theorem shows that we can compute an upper bound for $\mathbb{E}[J_{\text{robust}}(\bar{X}(0:t:N+1))]$ based on the J_{robust}. canonical form.

Theorem 6: For a given STL formula ψ, $\mathbb{E}[J_{\text{robust}}(\bar{X}(0:t:N+1))]$ can be upper bounded by

$$\max_{i\in\{1,\ldots,L\}}\min_{j\in\{1,\ldots,m_i\}}(\hat{\eta}_{ij} + \hat{d}_{ij}) + \kappa, \tag{8}$$

where $\hat{\eta}_{ij}$, $i\in\{1,\ldots,L\}, j\in\{1,\ldots,m_i\}$, are affine functions of $\bar{u}(0:t-1:N)$, $x(0)$, and $w(0),\ldots,w(t-1)$. The constants $\hat{d}_{ij}$ and κ are respectively a weighted sum of c_k,d_k and a function of a_k,b_k, $k\in\{1,\ldots,n\}$.

Proof: With focus on the canonical form (5), let $Y_{ij} = \eta_{ij} + \bar{W}(0:t:N)\lambda_{ij}$ with column vector $\lambda_{ij} := [\lambda_{ij}^0;\lambda_{ij}^1;\ldots;\lambda_{ij}^{N-1}]$, and $\lambda_{ij}^k\in\mathbb{R}^n$, $k\in\{0,1,\ldots,N-1\}$. Considering the support and moment interval of the components of $W(\tau),\tau\in\{t,\ldots,N-1\}$, each random variable Y_{ij} has the following support and moment interval (similar to (7))

$$I_{Y_{ij}} = [\hat{a}_{ij} + \hat{\eta}_{ij}, \hat{b}_{ij} + \hat{\eta}_{ij}],\ \mathbb{M}_{Y_{ij}} = [\hat{c}_{ij} + \hat{\eta}_{ij}, \hat{d}_{ij} + \hat{\eta}_{ij}] \tag{9}$$

where $\hat{\eta}_{ij} := \eta_{ij} + \sum_{k=0}^{t-1} w^T(k)\lambda_{ij}^k$, $[\hat{a}_{ij}, \hat{b}_{ij}] = \sum_{k=t}^{N-1} I_{W(k)}^T \lambda_{ij}^k$, and $[\hat{c}_{ij}, \hat{d}_{ij}] = \sum_{k=t}^{N-1} \mathbb{M}_{W(k)}^T \lambda_{ij}^k$. We utilize Lipschitz continuity of the min function to get

$$\min_j Y_{ij} \leq \min_j \left(\hat{\eta}_{ij} + \sum_{k=t}^{N-1} \mathbb{E}[W(k)]^T \lambda_{ij}^k \right) + \sum_{k=t}^{N-1} \max_j \|\lambda_{ij}^k\|_2 \|W(k) - \mathbb{E}[W(k)]\|_2.$$

By taking maximum and then expectation we have

$$\mathbb{E}[J_{\text{robust}}] \leq \max_i \min_j \left(\hat{\eta}_{ij} + \hat{d}_{ij}\right) + \sum_{k=t}^{N-1} \max_{i,j} \|\lambda_{ij}^k\|_2 \mathbb{E}\left[\|W(k) - \mathbb{E}[W(k)]\|_2\right] \quad (10)$$

The last term in (10) can also be upper bounded using Popoviciu's inequality on variances [4] as

$$\kappa := \sum_{k=t}^{N-1} \max_{i,j} \frac{1}{2} \|\lambda_{ij}^k\|_2 \left[\sum_{s=1}^{n} (b_s - a_s)^2\right]^{1/2}. \quad (11)$$

■

Hence, as we are minimizing the cost function in (4), we can replace $\mathbb{E}\left[J_{\text{robust}}(\bar{X}(0:t:N+1))\right]$ by $\max_i \min_j (\hat{\eta}_{ij} + \hat{d}_{ij})$ due to κ in (11) being a constant independent of the input. Note that the approximation methodology of Theorem 6 is applicable also to the min-max canonical form (6).

B. Normal distribution

The upper bound on the objective function in the previous section is not directly applicable to unbounded support disturbances. Here, we address disturbances with normal distribution separately due to their wide use in engineering applications.

Suppose that for any $t \in \mathbb{N}$, $W(t)$ is normally distributed with mean $\mathbb{E}[W(t)] = 0$ and covariance matrix $\Sigma_{W(t)}$, i.e., $W(t) \sim \mathcal{N}(0, \Sigma_{W(t)})$. The explicit form of $X(\tau)$ in (2) and the fact that normal distribution is closed under affine transformations result in normal distribution for $X(\tau)$, $\tau \in \mathbb{N}$. Its expected value and covariance matrix with an observed value $x(t)$ of $X(t)$ are $\mu_\tau = \Phi(\tau,t)x(t) + \sum_{k=t}^{\tau-1} \Phi(\tau,k+1)B(k)u(k)$ and $\Sigma_\tau = \sum_{k=t}^{\tau-1} \Phi(\tau,k+1)\Sigma_{W(k)}\Phi(\tau,k+1)^T$, $\quad \tau \geq t \geq 0$.

In this section we use the representation in Theorem 4, which states that J_{robust} can be written in either of the forms

$$\max_{i\in\{1,\ldots,L\}} \min_{j\in\{1,\ldots,m_i\}} Y_{ij} \quad \text{or} \quad \min_{i\in\{1,\ldots,K\}} \max_{j\in\{1,\ldots,n_i\}} Z_{ij} \quad (12)$$

with Y_{ij} and Z_{ij} being affine functions of disturbance, thus normally distributed random variables. With focus on these canonical representations for J_{robust}, we employ next theorem from [11] for computing an upper bound for $\mathbb{E}\left[J_{\text{robust}}\right]$ based on higher order moments of Y_{ij} and Z_{ij}.

Theorem 7: Considering the canonical forms in (12) for J_{robust} as a function of random variables Y_{ij} and Z_{ij}, $\mathbb{E}\left[J_{\text{robust}}\right]$ can be upper bounded by

$$\mathbb{E}\left[\max_{i\in\{1,\ldots,L\}} \min_{j\in\{1,\ldots,m_i\}} Y_{ij}\right] \leq \left(\sum_{i=1}^{L}\sum_{j=1}^{m_i} \mathbb{E}[Y_{ij}^p]\right)^{1/p}, \quad (13)$$

$$\mathbb{E}\left[\min_{i\in\{1,\ldots,K\}} \max_{j\in\{1,\ldots,n_i\}} Z_{ij}\right] \leq \min_{i\in\{1,\ldots,K\}} \left(\sum_{j=1}^{n_i} \mathbb{E}\left[Z_{ij}^p\right]\right)^{1/p}. \quad (14)$$

with $p > 0$ being an even integer.

Proof: The proof is based on the relation between the infinity norm and p-norm of a vector and Jensen's inequality. For brevity, we refer to Corollaries 7 and 8 in [11]. ■

Note that random variables Y_{ij} and Z_{ij} in (13)-(14) are normally distributed. Higher order moments of normal random variables can be computed analytically in a closed form as a function of the first two moments, i.e., mean and variance. More specifically, for a normal random variable Z with mean μ and variance σ^2, the p-th moment has a closed form as $\mathbb{E}\left[Z^p\right] = \sigma^p i^{-p} H_p(i\mu/\sigma)$ where i is the imaginary unit and

$$H_p(z) = p! \sum_{l=0}^{p/2} \frac{(-1)^l z^{p-2l}}{2^l l! (p-2l)!} \quad (15)$$

is the p-th Hermite polynomial [1, Chapter 22 and 26].

In the next section we discuss how to cope with the second challenge of characterizing the feasible set of the optimization restricted by the chance constraint (4c).

VII. Under Approximation of Chance Constraints

In this section, we discuss methods for computing conservative approximations of the chance constraints in (4c) as linear constraints. For the sake of compact notation, we indicate the stochastic process $\Xi(0:N) = X(0)X(1)\ldots X(N)$ only by Ξ_N without declaring its dependency on the state, input, and disturbance. Recall the chance constraint (4c) as $\Pr\left[\Xi_N \models \varphi\right] \geq 1-\delta$. In order to transform this constraint into linear inequalities, we first show in the following theorem, that this constraint can be transformed into similar probabilistic constraints on $(\Xi_N, \tau) \models \pi$, with π being an atomic predicate. Then in Sections VII-A and VII-B, we discuss how to transform the resulting constraints with atomic predicates into linear inequalities for the cases of arbitrary random variables with known bounded support and moment interval and of normally distributed random variables.

Theorem 8: For any formula φ and a constant $\vartheta \in (0,1)$, constraints of the forms

$$\Pr\left[(\Xi_N, t) \models \varphi\right] \geq \vartheta \text{ and } \Pr\left[(\Xi_N, t) \models \varphi\right] \leq \vartheta \quad (16)$$

can be transformed inductively on the structure of φ into similar probabilistic constraints on $(\Xi_N, \tau) \models \pi$, $\tau \geq t$, with π being an atomic predicate.

Proof: The proof is inductive on the structure of the formula φ as discussed in the following three cases.

Case I: $\varphi = \neg\varphi_1$ we have the following equivalences

$$\begin{aligned}
\Pr\left[(\Xi_N,t) \models \neg\varphi_1\right] \geq \vartheta &\Leftrightarrow \Pr\left[(\Xi_N,t) \not\models \varphi_1\right] \geq \vartheta \\
&\Leftrightarrow \Pr\left[(\Xi_N,t) \models \varphi_1\right] \leq 1-\vartheta, \\
\Pr\left[(\Xi_N,t) \models \neg\varphi_1\right] \leq \vartheta &\Leftrightarrow \Pr\left[(\Xi_N,t) \not\models \varphi_1\right] \leq \vartheta \\
&\Leftrightarrow \Pr\left[(\Xi_N,t) \models \varphi_1\right] \geq 1-\vartheta.
\end{aligned}$$

Case II: $\varphi = \varphi_1 \wedge \varphi_2$ we obtain the following inequalities by using the fact that for possibly joint events $\mathscr{A}$ and $\mathscr{B}$, it holds that $\Pr[\mathscr{A} \wedge \mathscr{B}] \geq \vartheta \Leftrightarrow \Pr(\neg\mathscr{A} \vee \neg\mathscr{B}) \leq 1-\vartheta$ and $\Pr(\mathscr{A} \vee \mathscr{B}) \leq \Pr[\mathscr{A}] + \Pr[\mathscr{B}]$.

$$\Pr[(\Xi_N,t) \models \varphi_1 \wedge \varphi_2] \geq \vartheta \Leftrightarrow \Pr\left[(\Xi_N,t) \models \varphi_1 \wedge (\Xi_N,t) \models \varphi_2\right] \geq \vartheta$$

$$\Leftrightarrow \Pr[(\Xi_N,t)\not\models \varphi_1 \vee (\Xi_N,t)\not\models \varphi_2] \le 1-\vartheta$$
$$\Leftarrow \Pr[(\Xi_N,t)\not\models \varphi_1] + \Pr[(\Xi_N,t)\not\models \varphi_2] \le 1-\vartheta$$
$$\Leftarrow \Pr[(\Xi_N,t)\not\models \varphi_i] \le \frac{1-\vartheta}{2} \quad i=1,2. \tag{17}$$

Now consider the second possibility:

$$\Pr[(\Xi_N,t)\models \varphi_1\wedge\varphi_2] \le \vartheta \Leftrightarrow \Pr[(\Xi_N,t)\models \neg\varphi_1\vee\neg\varphi_2] \ge 1-\vartheta$$
$$\Leftrightarrow \Pr[(\Xi_N,t)\models \neg\varphi_1 \vee (\varphi_1\wedge\neg\varphi_2)] \ge 1-\vartheta$$
$$\Leftrightarrow \Pr[(\Xi_N,t)\models \neg\varphi_1] + \Pr[(\Xi_N,t)\models \varphi_1\wedge\neg\varphi_2] \ge 1-\vartheta, \tag{18}$$

where the last line of (18) is due to the fact that the events are disjoint. Assuming that the probabilities of these two events are lower bounded by the same values, i.e., $(1-\vartheta)/2$, we have the inequalities

$$\Pr[(\Xi_N,t)\models\neg\varphi_1] \ge \frac{1-\vartheta}{2}, \quad \Pr[(\Xi_N,t)\models\varphi_1\wedge\neg\varphi_2] \ge \frac{1-\vartheta}{2},$$

which are in the form of inequalities discussed previously. Note that Equations (17) to (18) discuss the case of having conjunction of two STL formulas. The results can be easily extended to conjunction of n STL formulas by replacing $(1-\vartheta)/2$ with $(1-\vartheta)/n$.

Case III: $\varphi = \varphi_1 \mathscr{U}_{[a,b]} \varphi_2$ The satisfaction $(\Xi_N,t)\models \varphi_1 \mathscr{U}_{[a,b]}\varphi_2$ is equivalent to $\bigvee_{j=t+a}^{t+b} \psi_j$ with disjoint events

$$\psi_j = \bigwedge_{i=t}^{t+a-1} (\Xi_N,i)\models \varphi_1 \bigwedge_{i=a+t}^{j-1} (\Xi_N,i)\models (\varphi_1\wedge\neg\varphi_2)\wedge(\Xi_N,j)\models\varphi_2.$$

Thus $\Pr\left[(\Xi_N,t)\models \varphi_1\mathscr{U}_{[a,b]}\varphi_2\right] \ge \vartheta$ is equivalent to $\sum_{j=t+a}^{t+b}\Pr[\psi_j] \ge \vartheta$. Assuming the probabilities of events are lower bounded by the same values, we have $\Pr[\psi_j] \ge \vartheta/(b-a+1)$ for $j=a+t,\ldots,b+t$, which again can be reduced as in Case II.

The second possible probabilistic constraint in Case III can be obtained as

$$\Pr\left[(\Xi_N,t)\models \varphi_1\mathscr{U}_{[a,b]}\varphi_2\right] \le \vartheta \Leftrightarrow \Pr\left[\bigvee_{j=a+t}^{b+t}\psi_j\right] \le \vartheta$$
$$\Leftrightarrow \sum_{j=t+a}^{t+b}\Pr[\psi_j] \le \vartheta \Leftrightarrow \Pr[\psi_j] \ge \vartheta/(b-a+1), \tag{19}$$

which can be again reduced as in Case II. Here also, we used the fact that ψ_j consists of disjoint events and we assume that he probabilities of events are lower bounded by the same value, i.e., by $\vartheta/(b-a+1)$, for $j=a+t,\ldots,b+t$. ■

Remark 9: In order to reduce the level of conservatism, one might allow non-uniform risk allocation in Theorem 8. For instance, in the last line of (17), one can replace the two upper bounds $(1-\vartheta)/2$ with δ_1 and δ_2 such that $\delta_1+\delta_2 = 1-\vartheta$, and take them as part of the optimization. As we see later, these quantities will appear in the constraints through logarithm or the inverse of quantile functions. Then, the optimization problem in both cases will have nonlinear inequalities and its complexity depends on the number of variables, which results in larger computational complexity compare to uniform risk allocation and is not scalable specially due to the increasing number of δ_t as a function of length of STL formula.

So far, we have shown how to reduce the chance constraint (4c) inductively to inequalities of the form (16) with atomic predicates. In the rest of this section, we discuss their corresponding linear inequalities for the two types of probability distributions considered in this paper.

A. Arbitrary probability distributions with bounded support

To transform the chance constraints into linear constraints in the case of disturbances with arbitrary probability distributions, we apply an approximation method based on the upper bound proposed by [5]. Let Z_i, $i\in\{1,\ldots,n\}$, be a random variable with bounded support $[a_i,b_i]$ and expectation $\mathbb{E}[Z_i]$ belonging to the moment interval $\mathbb{M}_i$. Define $Z=\sum_{i=1}^n Z_i$ and $\mathbb{E}(Z)=\sum_{i=1}^n\mathbb{E}[Z_i]$. We derive an inequality for this generic random variable Z, which will be applied to $\alpha(X(t))$ in the sequel. Using Chernoff-Hoeffding inequality, the following upper bound exists

$$\Pr[Z-\mathbb{E}[Z]\le -\varsigma] \le \exp\left(\frac{-\varsigma^2}{\nu\sum_{i=1}^n(b_i-a_i)^2}\right), \ \forall\varsigma\ge 0. \tag{20}$$

where $\nu>0$ is a constant [16]. If $Z_1,\ldots,Z_n$ are dependent, then the inequality applies with a constant $\nu=\chi(\hat{G})/2$, where $\hat{G}$ denotes the indirected dependency graph of $Z_1,\ldots,Z_n$ and $\chi(\hat{G})$ is the chromatic number of the graph $\hat{G}$ defined as the minimum number of colors required to color $\hat{G}$. For the independent case, $\chi(\hat{G})=1$. The expression for the right tail probability is derived identically.

Consider the chance constraints (16) with atomic predicate $\varphi=\{\alpha\ge 0\}$, where $\alpha(x)=\alpha_0+\alpha_1^T x$ is an affine function evaluated at $X(\tau)$, $\tau\in\{t,t+1,\ldots,N\}$. Since $X(\tau)$ is a random variable with support and moment interval defined in (7), $\alpha(X(\tau))$ is itself a random variable with the following support and moment interval

$$I_{\alpha(X(\tau))} = [\tilde{a}_\tau+\tilde{C}_\tau, \tilde{b}_\tau+\tilde{C}_\tau],\ \mathbb{M}_{\alpha(X(\tau))} = [\tilde{c}_\tau+\tilde{C}_\tau, \tilde{d}_\tau+\tilde{C}_\tau] \tag{21}$$

where $\tilde{C}_\tau := \alpha_0+\alpha_1^T\bar{C}_\tau$ is an affine function of input variables, $[\tilde{a}_\tau,\tilde{b}_\tau] := \alpha_1^T[\bar{a}_\tau,\bar{b}_\tau]$, and $[\tilde{c}_\tau,\tilde{d}_\tau] := \alpha_1^T[\bar{c}_\tau,\bar{d}_\tau]$.

Applying (20) with $Z=\alpha(X(\tau))$ and $\varsigma=\mathbb{E}[Z]$, we obtain

$$\Pr[(\Xi_N,\tau)\models\varphi] \ge 1-\delta \Leftrightarrow \Pr[\alpha(X(\tau))>0] \ge 1-\delta$$
$$\Leftrightarrow \Pr[\alpha(X(\tau))\le 0] \le \delta \Leftarrow \exp\left(\frac{-\varsigma^2}{\nu \mathsf{s}_\alpha}\right) \le \delta \Leftrightarrow \frac{-\varsigma^2}{\nu\mathsf{s}_\alpha} \le \log(\delta)$$
$$\Leftrightarrow -\varsigma^2 \le \nu\log(\delta)\mathsf{s}_\alpha \Leftarrow \varsigma \ge \sqrt{-\nu\log(\delta)\mathsf{s}_\alpha}, \tag{22}$$

where $\mathsf{s}_\alpha := \sum_{k=t}^{\tau-1}\left[\alpha_1^T\Phi(\tau,k+1)|I_{W(k)}|\right]^2$ with $|I_{W(k)}|$ being the length of $I_{W(k)}$. Hence, we can replace ς in (22) by the lower bound of its moment interval in (21), i.e., with $\tilde{c}_\tau+\tilde{C}_\tau$, which is a linear expression in input variables. Consequently, the chance constraint in (4) can be conservatively replaced by inequalities of the form

$$\tilde{c}_\tau+\tilde{C}_\tau \ge \sqrt{-\nu\log(\delta)\cdot\mathsf{s}_\alpha}. \tag{23}$$

For the second type of probabilistic inequalities in (16), we can again use (20) for the right tail probability; hence we have

$$\Pr[(\Xi_N,\tau)\models\varphi] \le 1-\delta \Leftarrow \Pr[\alpha(X(\tau))\ge 0] \le 1-\delta$$

$$\Leftarrow \exp\left(\frac{-\varsigma^2}{\nu s_\alpha}\right) \le 1-\delta, \tag{24}$$

and then following the same steps as in (22), we obtain the same linear expression for the chance constant as in (23) by only replacing δ by $1-\delta$ in the related expressions.

By replacing the expectation of the objective function with its upper bound given in Theorem 6, and by substituting probabilistic constraints with their linear approximations, optimization problem (4) can be then recast as a mixed integer linear programming (MILP) problem. This is due the presence of nonlinear functions only in the form of min, max, and absolute value, which all can be expressed in terms of Boolean variables and linear functions. The resulting optimization can then be solved using the available MILP solvers [2], [19].

B. Normal distribution

To transform the chance constraints into linear constraints in the case of having normally distributed random variables, we use the quantile of the normal distribution. By definition, for a normal random variable X with mean μ and variance σ^2,

$$\Pr[X \le b] \le \delta \Leftrightarrow F^{-1}(\delta) \ge b \Leftrightarrow \mu + \sigma\phi^{-1}(\delta) \ge b \tag{25}$$

$$\Pr[X \le b] \ge \delta \Leftrightarrow F^{-1}(\delta) \le b \Leftrightarrow \mu + \sigma\phi^{-1}(\delta) \le b \tag{26}$$

where F^{-1} denotes the inverse of the cumulative distribution function or the quantile function and ϕ^{-1} is the inverse of the error function of a normally distributed random variable.

Recall the chance constraints (16) with $\varphi = \{\alpha \ge 0\}$. Since $\alpha(\cdot)$ is an affine function of normally distributed state variables, it is also normal with appropriately defined mean μ_τ and variance σ_τ^2. Hence, we can directly use (25)-(26) as

$$\Pr[(\Xi_N,\tau) \models \varphi] \ge 1-\delta \Leftrightarrow \Pr[\alpha(X(\tau)) > 0] \ge 1-\delta \tag{27}$$
$$\Leftrightarrow \Pr[\alpha(X(\tau)) \le 0] \le \delta \Leftrightarrow F^{-1}(\delta) \ge 0 \Leftrightarrow \mu_\tau + \sigma_\tau\phi^{-1}(\delta) \ge 0,$$

$$\Pr[(\Xi_N,\tau) \models \varphi] \le 1-\delta \Leftrightarrow \Pr[\alpha(X(\tau)) > 0] \le 1-\delta \tag{28}$$
$$\Leftrightarrow \Pr[\alpha(X(\tau)) \le 0] \ge \delta \Leftrightarrow F^{-1}(\delta) \le 0 \Leftrightarrow \mu_\tau + \sigma_\tau\phi^{-1}(\delta) \le 0.$$

Therefore, the chance constraint can be replaced by the equivalent linear constraint (27) or (28), depending on the type of the constraint we have.

By replacing the expectation of the objective function by its upper bound given in Theorem 7, and by substituting chance constraints with their linear approximations, optimization problem (4) can be recast as a (possibly convex) nonlinear optimization problem with linear constraints, which can be solved using algorithms such as interior point method [23] or multi-start sequential quadratic programming (SQP) [22].

VIII. Experimental Results

We use our synthesis approach for controlling the temperature in a building. The thermal model of the building is: $X(t+1) = AX(t) + Bu(t) + W(t)$; where $X \in \mathbb{R}^n$ is the temperatures of walls and rooms and input $u \in \mathbb{R}^m$ includes the air mass flow rate and discharge air temperature of conditioned air into each thermal zone. Matrices A, B are obtained after linearizing and discretizing the model presented in [15], [24] with sampling time $t_s = 30$ minutes. Disturbance $W(\cdot)$ aggregates various unmodeled dynamics of the system [15]. We control the temperature of one room in the building, which is the last element of state X denoted by X_5, with $n = 5$ and $m = 1$. Unlike [24] that considers deterministic disturbances, we assume stochastic disturbances with a reference $w_r(t)$ and perturbed by uniformly distributed random vectors $e(t)$ with support $[-1.5, 1.5]^n$, i.e., $W(t) = w_r(t) + e(t)$.

Temperature dynamics depend also on room occupancy indicated by a known signal $\mathrm{occ} : \mathbb{N} \to \{-1, 1\}$, where $\mathrm{occ}(t) = 1$ if the room is occupied at time t and $\mathrm{occ}(t) = -1$ otherwise. We are interested in keeping the room temperature in the interval $[T_r - \Delta, T_r + \Delta]$ whenever the room is occupied. The reference temperature is $T_r = 68°\mathrm{F}$ and $\Delta = 1$ is the acceptable variation. For this to happen, we allow the controller to change the temperature within 3 time steps, i.e., $[0, 2]$, when the room is occupied. Temperature should also always stay in $[66, 72]$. This desired behavior can be expressed via the STL specification $\varphi = \psi_1 \wedge \psi_2$, where $\psi_1 = \Box_{[0,N]}(66 \le X_5 \le 72)$ and $\psi_2 = \Box_{[0,N]}\Big(\mathrm{occ} = 1 \to \Diamond_{[0,2]}\big((|X_5 - T_r| \le \Delta)\,\mathscr{U}_{[0,N]}(\mathrm{occ} = -1)\big)\Big)$. In optimization problem (4), we consider the chance constraint (4c) defined with the specification φ. We choose the objective function (4a) as $\mathbb{E}[-\rho^{\psi_2}(\bar{X}(0:t:N), 0)] + \gamma_u \sum_{k=0}^{N-1} ||u(k)||_1$, which includes robustness of ψ_2. Hence, the optimization tries to satisfy both ψ_1 and ψ_2 with probability $1-\delta$, but puts more emphasis on ψ_2 by maximizing its robustness in a tradeoff with the consumed energy weighted by a constant $\gamma_u > 0$. We approximate $\mathbb{E}[-\rho^{\psi}(\tilde{X}(0:t:N), 0)]$ using (8) and transform the chance constraint into linear inequalities using the approach of Section VII-A. We also assume that inputs are bounded, i.e., $u(t) \in U = [0, 300]$. We control the room temperature for 23 hours ($N = 46$) and select $\delta = 0.1$ so that the obtained control input provides 90% confidence on satisfaction of the desired behavior.

We perform 200 simulations using MATLAB R2016b on a 3.1 GHz Intel Core i5 processor. Figure 1 shows the results of these simulations. The top plot shows the occupancy signal. The middle plot illustrates the average, minimum, and maximum of the obtained room temperatures over 200 simulations as a function of time. It also shows the minimum and maximum room temperature bounds in Fahrenheit. The controller ensures that the room temperature enters the desired interval within two time steps once the occupancy signal is one and stays there as long as the room is occupied. We witnessed that our proposed over approximation of the chance constraints is infeasible in 6.5% of the simulations, but including the robustness in the objective function pushes the closed-loop system towards satisfying the specification in all cases (cf. Remark 3). The bottom plot shows the average, minimum, and maximum of the air flow rate in $\left[\frac{\mathrm{ft}^3}{\mathrm{min}}\right]$, which indicates that the input constraint is not violated.

Note that assessing the level of conservatism in replacing the objective function (4a) with an upper bound using (8) is analytically cumbersome. However, for this case study, we have calculated the objective function using Monte Carlo simulation and compared its values against the upper bound (8).

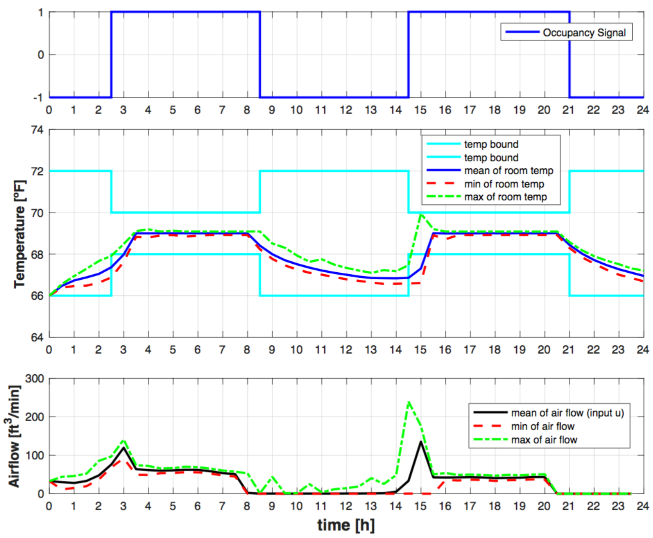

Fig. 1. Controlling the room temperature using SHMPC in the presence of normally distributed disturbance and STL constraints.

TABLE I
COMPARISON OF THE STATISTICS OF THE FAN ENERGY CONSUMPTION USING RMPC AND SHMPC APPROACHES.

Computational Methods	Fan energy consumption [kWh]	Average computation time [s]
RMPC	$\mu_1 = 11.3325$, $\sigma_1 = 0.0346\mu_1$	80.2693
SHMPC	$\mu_2 = 9.2784$, $\sigma_2 = 0.17211\mu_2$	15.3630

The comparison shows that the average and maximum relative errors are respectively 15.22% and 15.54% in 200 simulations.

To further illustrate the performance of our method, we compare our SHMPC approach with the robust MPC (RMPC) approach of [24]. Table VIII shows total fan energy consumption, which is proportional to the cubic of airflow, and the computation times for both approaches. For RMPC and SHMPC, we report the average computation time and the average and standard deviation of the total energy consumption using the sum of cubes of the optimal input sequences corresponding to 200 simulations. Since RMPC is more conservative compared to SHMPC, the average energy consumption is higher for the RMPC controller compared to the SHMPC controller: the SHMPC controller achieves 18% reduction of total energy consumption on average compared to RMPC.

IX. CONCLUSIONS

In this paper, we presented shrinking horizon model predictive control (SHMPC) for stochastic linear systems with constraints encoded as signal temporal logic (STL) specifications. The goal of SHMPC is to obtain an optimal control sequence that guarantees satisfaction of STL specifications with a probability greater than a given threshold. We provided an approximation technique that gives an upper bound on the objective function and conservatively replaces chance constraint with linear inequalities. Our approximation relies on knowing only the support and moment intervals of disturbance. We also discussed how the approximation can be customized for normal disturbances.

REFERENCES

[1] M. A. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. National Bureau of Standards, US Government Printing Office, Washington DC, 1964.

[2] A. Atamtürk and M. W. P. Savelsbergh. Integer-programming software systems. *Annals of Operations Research*, 140(1):67–124, 2005.

[3] A. Bemporad, W. P. M. H. Heemels, and B. De Schutter. On hybrid systems and closed-loop MPC systems. *IEEE Transactions on Automatic Control*, 47(5):863–869, May 2002.

[4] Rajendra Bhatia and Chandler Davis. A better bound on the variance. *The American Mathematical Monthly*, 107(4):353–357, 2000.

[5] O. Bouissou, E. Goubault, S. Putot, A. Chakarov, and S. Sankaranarayanan. Uncertainty propagation using probabilistic affine forms & concentration of measure inequalities. In *TACAS*, LNCS 9636. Springer, 2016.

[6] A. Carvalho, Y. Gao, S. Lefevre, and F. Borrelli. Stochastic predictive control of autonomous vehicles in uncertain environments. In *AVEC*, 2014.

[7] P. J. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Academic Press, New York, 2nd edition, 1984.

[8] B. De Schutter and T. J. J. van den Boom. MPC for continuous piecewise-affine systems. *Syst. & Contr. Letters*, 52(3–4):179–192, 2004.

[9] A. Donzé and O. Maler. Robust satisfaction of temporal logic over real-valued signals. In *FORMATS*, LNCS 6246. Springer, 2010.

[10] G. E. Fainekos and G. J. Pappas. Robustness of temporal logic specifications for continuous-time signals. *Theor. Comput. Sci.*, 410(42):4262–4291, 2009.

[11] S. S. Farahani, R. Majumdar, V.S. Prabhu, and S. Soudjani. Shrinking horizon model predictive control with chance-constrained signal temporal logic specifications. In *ACC*. IEEE, 2017.

[12] S. S. Farahani, V. Raman, and R. M. Murray. Robust model predictive control for signal temporal logic synthesis. In *ADHS*. IFAC, 2015.

[13] S. S. Farahani, T. van den Boom, H. van der Weide, and B. De Schutter. An approximation method for computing the expected value of max-affine expressions. *European Journal of Control*, 27:17–27, 2016.

[14] J. M. Grosso, C. Ocampo-Martinez, V. Puig, and B. Joseph. Chance-constrained model predictive control for drinking waternetworks. *Journal of process control*, 24:504–516, 2014.

[15] M. Maasoumy Haghighi. *Controlling Energy Efficient Buildings in the Context of Smart Grid: A Cyber Physical System Approach*. PhD thesis, University of California, Berkeley, Berkeley, California, December 2013.

[16] S. Janson. Large deviations for sums of partly dependent random variables. *Random Structures Algorithms*, 24(3):234–248, 2004.

[17] S. Jha and V. Raman. Automated synthesis of safe autonomous vehicle control under perception uncertainty. In *NFM*, LNCS 9690. Springer, 2016.

[18] X. Jin, A. Donzé, J. V. Deshmukh, and S. A. Seshia. Mining requirements from closed-loop control models. *IEEE Trans. on CAD of Integrated Circuits and Systems*, 34(11):1704–1717, 2015.

[19] J. T. Linderoth and T. K. Ralphs. Noncommercial software for mixed-integer linear programming. In *Reinforcement Learning: State-Of-The-Art*, CRC Press Operations Research Series, pages 253–303. 2005.

[20] J. M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, 2002.

[21] O. Maler and D. Nickovic. Monitoring temporal properties of continuous signals. In *FORMATS/FTRTFT*, LNCS 3253. Springer, 2004.

[22] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, 1996.

[23] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, Pennsylvania, 1994.

[24] V. Raman, A. Donzé, D. Sadigh, R. M. Murray, and S. A. Seshia. Reactive synthesis from signal temporal logic specifications. In *HSCC*. ACM, 2015.

[25] D. Sadigh and A. Kapoor. Safe control under uncertainty with probabilistic signal temporal logic. In *RSS*, 2016.

[26] S. Sadraddini and C. Belta. Robust temporal logic model predictive control. In *Allerton Conference*, 2015.

[27] A. T. Schwarm and M. Nikolaou. Chance-constrained model predictive control. *AIChE Journal*, 45(8):1743–1752, 1999.

[28] M. P. Vitus. *Stochastic Control Via Chance Constrained Optimization and its Application to Unmanned Aerial Vehicles*. PhD thesis, Stanford University, Stanford, California, March 2012.

[29] M. P. Vitus and C. J. Tomlin. A hybrid method for chance constrained control in uncertain environments. In *CDC*. IEEE, 2012.